

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

**Amendments to the Claims:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1. (Original) A method of creating a searchable archive accessible by a data processing system, comprising:
  - generating a domain structure and tokenized data from an archive data set, the domain structure including tokens corresponding to unique values in the archive data set and the tokenized data including token columns corresponding to value columns in the archive data set;
  - determining archive metadata from the domain structure and the tokenized data;
  - dividing the tokenized data into one or more token column segments;
  - determining token column segment metadata from the one or more token column segments;
  - creating one or more compressed token column segments from the token column segments;
  - creating one or more compacted files from the one or more compressed token column segments and the token column segment metadata; and
  - storing the one or more compacted files in a file system coupled to the data processing system.
2. (Original) The method of claim 1, wherein determining metadata further comprises determining a maximum value and a minimum value for each of the token columns.
3. (Original) The method of claim 1, wherein determining metadata further comprises determining a maximum tupleid and a minimum tupleid for each of the one or more token column segments.
4. (Original) The method of claim 1, further comprising:
  - dividing the domain structure into one or more domain structure segments;

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

determining metadata from the domain structure segments;  
compressing the one or more domain structure segments; and  
creating one or more compacted files further includes storing the compressed domain structure segments in the compacted file.

5. (Withdrawn) A method of retrieving a datum from a searchable archive by a data processing system, the searchable archive comprising a metadata file and one or more compacted files, comprising:

selecting a selected compacted file from the one or more compacted files that may include the datum using the metadata file;  
accessing the selected compacted file;  
selecting a selected compressed segment from one or more compressed segments in the selected compacted file using metadata stored in the compacted file;  
generating a decompressed segment from the selected compressed segment; and  
searching the decompressed segment to determine if the decompressed segment includes the datum.

6. (Withdrawn) The method of claim 5 wherein:

selecting a selected compacted file is performed by a search process; and  
accessing the selected compacted file, selecting a selected compressed segment, generating a decompressed segment, and searching the decompressed segment are performed by one or more search agents invoked by the search process.

7. (Original) A method of creating a searchable archive accessible by a data processing system, comprising:

generating a domain structure and tokenized data from archive data;  
determining metadata from the tokenized data;  
generating a set of bit vectors from the tokenized data;  
creating one or more compacted files from the set of bit vectors; and

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

storing the one or more compacted files in a file system coupled to the data processing system.

8. (Original) The method of claim 7, wherein the tokenized data set includes one or more columns of tokens and extracting archive metadata further comprises determining a maximum token value and a minimum token value for each of the one or more columns of tokens.

9. (Withdrawn) A method of retrieving a datum from a searchable archive by a data processing system, the searchable archive comprising a metadata file and one or more compacted files, comprising:

- selecting a selected compacted file from the one or more compacted files that may include the datum using the metadata;
- accessing the selected compacted file;
- selecting one or more bit vectors from the selected compacted file; and
- performing a Boolean operation on the bit vectors included in the selected compacted file to determine if the datum is stored in the selected compacted file.

10. (Withdrawn) The method of claim 9, wherein:

- selecting a selected compacted file is performed by a search process; and
- accessing the selected compacted file and performing a Boolean operation is performed by one or more search agents invoked by the search process.

11. (Original) A data processing system for creating a searchable archive, comprising:  
a processor; and

a memory coupled to the processor, the memory having program instructions executable by the processor stored therein, the program instructions including:

- generating a domain structure and tokenized data from an archive data set, the domain structure including tokens corresponding to unique values in the archive data set and the tokenized data including token columns corresponding to value columns in the archive data set;

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

determining archive metadata from the domain structure and the tokenized data;  
dividing the tokenized data into one or more token column segments;  
determining token column segment metadata from the one or more token column segments;  
creating one or more compressed token column segments from the token column segments;  
creating one or more compacted files from the one or more compressed token column segments and the token column segment metadata; and  
storing the one or more compacted files in a file system coupled to the data processing system.

12. (Original) The data processing system of claim 11, the program instructions for determining metadata further including determining a maximum value and a minimum value for each of the token columns.

13. (Original) The data processing system of claim 11, the program instructions for determining metadata further including determining a maximum tupleid and a minimum tupleid for each of the one or more token column segments.

14. (Original) The data processing system of claim 11, the program instructions further including:

dividing the domain structure into one or more domain structure segments;  
determining metadata from the domain structure segments;  
compressing the one or more domain structure segments; and  
creating one or more compacted files further includes storing the compressed domain structure segments in the compacted file.

15. (Withdrawn) A data processing system for retrieving a datum from a searchable archive, the searchable archive comprising a metadata file and one or more compacted files, comprising:  
a processor; and

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

a memory coupled to the processor, the memory having program instructions executable by the processor stored therein, the program instructions including:

selecting a selected compacted file from the one or more compacted files that may include the datum using the metadata file;

accessing the selected compacted file;

selecting a selected compressed segment from one or more compressed segments in the selected compacted file using metadata stored in the compacted file;

generating a decompressed segment from the selected compressed segment; and

searching the decompressed segment to determine if the decompressed segment includes the datum.

16. (Withdrawn) The data processing system of claim 15, the program instructions further including:

selecting a selected compacted file is performed by a search process; and

accessing the selected compacted file, selecting a selected compressed segment, generating a decompressed segment, and searching the decompressed segment are performed by one or more search agents invoked by the search process.

17. (Original) A data processing system for creating a searchable archive, comprising:  
a processor; and

a memory coupled to the processor, the memory having program instructions executable by the processor stored therein, the program instructions including:

generating a domain structure and tokenized data from archive data;

determining metadata from the tokenized data;

generating a set of bit vectors from the tokenized data;

creating one or more compacted files from the set of bit vectors; and

storing the one or more compacted files in a file system coupled to the data processing system.

18. (Original) The data processing system of claim 17, wherein the tokenized data set

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

includes one or more columns of tokens, the program instructions for extracting archive metadata further including determining a maximum token value and a minimum token value for each of the one or more columns of tokens.

19. (Withdrawn) A data processing system for retrieving a datum from a searchable archive, the searchable archive comprising a metadata file and one or more compacted files, comprising:

a processor; and

a memory coupled to the processor, the memory having program instructions executable by the processor stored therein, the program instructions including:

selecting a selected compacted file from the one or more compacted files that may include the datum using the metadata;

accessing the selected compacted file;

selecting one or more bit vectors from the selected compacted file; and

performing a Boolean operation on the bit vectors included in the selected compacted file to determine if the datum is stored in the selected compacted file.

20. (Withdrawn) The data processing system of claim 19, wherein:

selecting a selected compacted file is performed by a search process; and

accessing the selected compacted file and performing a Boolean operation is performed by one or more search agents invoked by the search process.

21. (Original) A method of utilizing a searchable archive by a data processing system, comprising:

generating a domain structure and tokenized data from archive data;

determining archive metadata from the tokenized data;

dividing the tokenized data into one or more segments;

determining segment metadata from the one or more segments;

creating one or more compressed segments from the segments;

creating one or more compacted files from the one or more compressed segments and the segment metadata; and

**Appln No. 10/783,643**  
**Amdt date November 20, 2006**  
**Reply to Office action of October 20, 2006**

storing the one or more compacted files in a file system coupled to the data processing system.

22. (Original) The method of claim 21, further comprising:  
selecting a selected compacted file from the one or more compacted files that may include a datum using the archive metadata;  
accessing the selected compacted file;  
selecting a selected compressed segment from the one or more compressed segments in the selected compacted file using the segment metadata;  
generating a decompressed segment from the selected compressed segment; and  
searching the decompressed segment to determine if the decompressed segment includes the datum.
23. (Original) The method of claim 22 wherein:  
selecting a selected compacted file is performed by a search process; and  
accessing the selected compacted file, selecting a selected compressed segment, generating a decompressed segment, and searching the decompressed segment are performed by one or more search agents invoked by the search process.
24. (Original) A method of utilizing a searchable archive by a data processing system, comprising:  
generating a domain structure and tokenized data from archive data;  
determining archive metadata from the tokenized data;  
generating a set of bit vectors from the tokenized data;  
creating one or more compacted files from the set of bit vectors; and  
storing the one or more compacted files in a file system coupled to the data processing system.
25. (Original) The method of claim 24, further comprising:  
selecting a selected compacted file from the one or more compacted files that may include a datum using the archive metadata;  
accessing the selected compacted file;

selecting one or more bit vectors from the selected compacted file; and  
performing a Boolean operation on the bit vectors included in the to determine if  
the datum is stored in the compacted file.

26. (Original) The method of claim 25, wherein:  
selecting a selected compacted file is performed by a search process; and  
accessing the selected compacted file and  
performing a Boolean operation is performed by one or more search agents  
invoked by the search process.
27. (Original) A data processing system for utilizing a searchable archive, comprising:  
a processor; and  
a memory coupled to the processor, the memory having program instructions  
executable by the processor stored therein, the program instructions including:  
generating a domain structure and tokenized data from archive data;  
determining archive metadata from the tokenized data;  
dividing the tokenized data into one or more segments;  
determining segment metadata from the one or more segments;  
creating one or more compressed segments from the segments;  
creating one or more compacted files from the one or more compressed  
segments and the segment metadata; and  
storing the one or more compacted files in a file system coupled to the  
data processing system.
28. (Original) The data processing system of claim 27, the program instructions further  
including:  
selecting a selected compacted file from the one or more compacted files that may  
include a datum using the archive metadata;  
accessing the selected compacted file;  
selecting a selected compressed segment from the one or more compressed  
segments in the selected compacted file using the segment metadata;  
generating a decompressed segment from the selected compressed segment; and



searching the decompressed segment to determine if the decompressed segment includes the datum.

29. (Original) The data processing system of claim 28, wherein  
selecting a selected compacted file is performed by a search process; and  
accessing the selected compacted file, selecting a selected compressed segment, generating a decompressed segment, and searching the decompressed segment are performed by one or more search agents invoked by the search process.
30. (Original) A data processing system for utilizing a searchable archive, comprising:  
a processor; and  
a memory coupled to the processor, the memory having program instructions executable by the processor stored therein, the program instructions including:  
generating a domain structure and tokenized data from archive data;  
determining archive metadata from the tokenized data;  
generating a set of bit vectors from the tokenized data;  
creating one or more compacted files from the set of bit vectors; and  
storing the one or more compacted files in a file system coupled to the data processing system.
31. (Original) The data processing system of claim 30, the program instructions further including:  
selecting a selected compacted file from the one or more compacted files that may include a datum using the archive metadata;  
accessing the selected compacted file;  
selecting one or more bit vectors from the selected compacted file; and  
performing a Boolean operation on the bit vectors included in the to determine if the datum is stored in the compacted file.
32. (Original) The data processing system of claim 31, wherein:  
selecting a selected compacted file is performed by a search process; and  
accessing the selected compacted file and

**Appln No. 10/783,643**

**Amdt date November 20, 2006**

**Reply to Office action of October 20, 2006**

performing a Boolean operation is performed by one or more search agents invoked by the search process.